Probability Comprehensive Exam
January 15, 2016

Student Number: $\boxed{}$

Instructions: Complete up to 5 of the 8 problems, and **circle** their numbers below – the uncircled problems will **not** be graded.

1 2 3 4 5 6 7 8

Write **only on the front side** of the solution pages. A **complete solution** of a problem is preferable to partial progress on several problems.

1. Let X have $\mathbb{E}X = 0$ and Var $X = \sigma^2 > 0$. Show that if $c > 0$, then
$$\mathbb{P}(X > c) \leq \frac{\sigma^2}{\sigma^2 + c^2}.$$
Hint: write $c - X = (c - X)_+ - (c - X)_-$ and use Cauchy-Schwarz type arguments.

Solution: Note that
$$c = \mathbb{E}(c - X) \leq \mathbb{E}(c - X)_+ = \mathbb{E}(c - X)\mathbf{1}_{\{X<c\}}.$$
Using Cauchy-Schwarz, the right side is bounded by
$$\sqrt{\mathbb{E}(c - X)^2\mathbb{P}(X < c)} = \sqrt{(c^2 + \sigma^2)\mathbb{P}(X < c)}.$$
Therefore
$$\mathbb{P}(X < c) \geq \frac{c^2}{c^2 + \sigma^2}$$
and the result follows.

2. Let $X = (X_1, X_2)$ be a Gaussian vector with zero mean and covariance matrix
$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$
where $|\rho| < 1$. Find a matrix A such that $X = AZ$, where Z is a standard normal vector and derive the characteristic function of X as a function of ρ.

Solution: If $A : \mathbb{R}^2 \to \mathbb{R}^2$ is linear and we set $Y = AZ$, then we can compute the covariance matrix Σ of Y as
$$\Sigma_{i,j} = \mathbb{E}Y_iY_j = \sum_{k,l} \mathbb{E}A_{i,k}Z_kA_{j,l}Z_l = \sum_k A_{i,k}A_{j,k} = (AA^T)_{i,j}.$$
Therefore to find A in the statement of the problem, we need to solve $AA^T = \Sigma$. Putting $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we want
$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix}.$$
If we set $c = 0$ then this system becomes
$$\begin{aligned} a^2 + b^2 &= 1 \\ bd &= \rho \\ d^2 &= 1 \end{aligned}$$

If we set $d = 1, b = \rho$ and $a = \sqrt{1-\rho^2}$, then we obtain a solution.

To compute the characteristic function, we write

$$\mathbb{E}e^{it\cdot AZ} = \mathbb{E}e^{i\sum_j t_j (AZ)_j} = \mathbb{E}\exp\left(i\sum_{j,k} t_j A_{j,k} Z_k\right).$$

Now use independence of the Z_k's to obtain

$$\prod_k \mathbb{E}\exp\left(i\sum_j t_j A_{j,k} Z_k\right) = \prod_k \mathbb{E}\exp\left(i(A^T t)_k Z_k\right).$$

The inner term is the characteristic function of a standard normal random variable, evaluated at $(tA)_k$, so we obtain

$$\prod_k \exp\left(-\frac{1}{2}(A^T t)_k^2\right) = \exp\left(-\frac{1}{2}\|A^T t\|^2\right) = \exp\left(-\frac{1}{2}\langle AA^T t, t\rangle\right) = \exp\left(-\frac{1}{2}\langle \Sigma t, t\rangle\right).$$

In terms of ρ, this becomes

$$\phi_X(t_1, t_2) = \exp\left(-\frac{1}{2}(t_1^2 - 2\rho t_1 t_2 + t_2^2)\right).$$

3. Let $X_1, X_2, \ldots$ be i.i.d. uniform $(0,1)$ random variables. Show that

$$(X_1 \cdot \cdots \cdot X_n)^{1/n}$$

converges almost surely as $n \to \infty$ and compute the limit.

Solution: Let $Y_i = \log X_i$. Then if we write P_n for the above expression, one has

$$\log P_n = \frac{Y_1 + \cdots + Y_n}{n}.$$

If we can use the strong law of large numbers, then we will obtain

$$\log P_n \to \mathbb{E}Y_1, \text{ or } P_n \to e^{\mathbb{E}Y_1} \text{ almost surely.}$$

So we set to compute $\mathbb{E}Y_1$. Note that $Y_1 \le 0$ almost surely, so for $y \le 0$,

$$\mathbb{P}(Y_1 \le y) = \mathbb{E}(X_1 \le e^y) = e^y,$$

since X_1 is uniformly distributed. This means that $-Y_1$ is continuously distributed, nonnegative, and with $\mathbb{P}(-Y_1 \geq y) = \mathbb{P}(Y_1 \leq -y) = e^{-y}$ for $y \geq 0$. Using the tail-sum formula for expectation,

$$\mathbb{E}(-Y_1) = \int_0^\infty \mathbb{P}(-Y_1 \geq y)\ \mathrm{d}y = \int_0^\infty e^{-y}\ \mathrm{d}y = 1.$$

Therefore $\mathbb{E}Y_1 = -1$. We conclude that $\mathbb{E}|Y_1| = \mathbb{E}(-Y_1) = 1$ exists, so the strong law of large numbers applies and we obtain

$$(X_1 \cdot \dots \cdot X_n)^{1/n} \to \exp(\mathbb{E}Y_1) = e^{-1}.$$

4. Let $X_1, X_2, \dots$ be i.i.d. exponential variables with parameter 1 and set

$$M_n = \max\{X_1, \dots, X_n\}.$$

Find sequences (a_n) and (b_n) of real numbers such that $(M_n - a_n)/b_n$ converges in distribution.

Solution: For $x \in \mathbb{R}$, by the i.i.d. assumption,

$$\mathbb{P}(M_n \leq x) = \mathbb{P}(X_i \leq x \text{ for all } i = 1, \dots, n) = \mathbb{P}(X_1 \leq x)^n.$$

For $x \leq 0$, this is 0. For $x > 0$, one has

$$\mathbb{P}(X_1 \leq x) = \int_0^x e^{-t}\ \mathrm{d}t = 1 - e^{-x}.$$

Therefore

$$\mathbb{P}(M_n \leq x) = \begin{cases} 0 & \text{if } x \leq 0 \\ (1 - e^{-x})^n & \text{if } x > 0 \end{cases}.$$

Plugging in $x = \log n + c$, one obtains for n large

$$\mathbb{P}(M_n \leq \log n + c) = \left(1 - e^{-c-\log n}\right)^n = \left(1 - \frac{e^{-c}}{n}\right)^n \to e^{-e^{-c}}.$$

Note that $\lim_{c\to-\infty} e^{-e^{-c}} = 0$ and $\lim_{c\to\infty} e^{-e^{-c}} = 1$, and so since it is continuous (in particular right-continuous) the function $c \mapsto e^{-e^{-c}}$ is the distribution function for a probability measure. Therefore if we set Z to be a random variable with this distribution, and $a_n = \log n,\ b_n = 1$, one has

$$\mathbb{P}\left(\frac{M_n - a_n}{b_n} \leq c\right) = \mathbb{P}(M_n \leq \log n + c) \to e^{-e^{-c}},$$

and so $(M_n - a_n)/b_n \Rightarrow Z$.

5. Let $(N_t)_{t\geq 0}$ be a rate-λ Poisson process. Let $X_1, X_2, \ldots$ be i.i.d. random variables with $\mathbb{E}|X_1| < \infty$ (independent of the Poisson process as well) and define

$$S_t = \sum_{i=1}^{N_t} X_i.$$

Show that S_t/t converges in probability to a constant and compute this constant.

Solution: We can compute the characteristic function of S_t in terms of the characteristic function ϕ of X_1:

$$\begin{aligned}
\mathbb{E}e^{isS_t} = \sum_{n=0}^{\infty} \mathbb{E}e^{isS_t}\mathbf{1}_{\{N_t=n\}} &= \sum_{n=0}^{\infty} \mathbb{E}e^{is(X_1+\cdots+X_n)}\mathbf{1}_{\{N_t=n\}} \\
&= \sum_{n=0}^{\infty} \left(\mathbb{E}e^{isX_1}\right)^n \mathbb{P}(N_t = n) \\
&= e^{-\lambda t} \sum_{n=0}^{\infty} \phi^n(s) \frac{(\lambda t)^n}{n!} \\
&= e^{-\lambda t} \exp\left(\phi(s)\lambda t\right) \\
&= \exp\left(\lambda t(\phi(s) - 1)\right).
\end{aligned}$$

So the characteristic function of S_t/t is

$$\exp\left(\lambda t(\phi(s/t) - 1)\right) = \exp\left(\lambda s \left(\frac{\phi(s/t) - \phi(0)}{s/t}\right)\right).$$

As $t \to \infty$,

$$\frac{\phi(s/t) - \phi(0)}{s/t} \to \phi'(0) = \frac{\mathrm{d}}{\mathrm{d}a}\mathbb{E}e^{iaX_1} = i\mathbb{E}X_1.$$

So for each s,

$$\mathbb{E}e^{is(S_t/t)} \to e^{i\lambda s\mathbb{E}X_1},$$

which is the characteristic function of the constant variable $\lambda\mathbb{E}X_1$. By the continuity theorem, one has

$$S_t/t \Rightarrow \lambda\mathbb{E}X_1.$$

Since convergence in distribution to a constant implies convergence in probability, $S_t/t \to \lambda\mathbb{E}X_1$ in probability.

6. Let $X_1, X_2, \ldots$ be i.i.d. standard normal random variables and for $x \in (-1, 1)$, set

$$Y = \sum_{n=1}^{\infty} x^n X_n.$$

Show that the sum defining Y converges and find its distribution.

Solution: To show that the sum converges, we can compute:

$$\mathbb{P}(|X_n| > n) = \frac{2}{\sqrt{2\pi}} \int_n^\infty e^{-t^2/2}\ \mathrm{d}t \le Ce^{-n^2/2}.$$

(Here we are using the approximation $\mathbb{P}(X_n > x) \sim \frac{e^{-x^2/2}}{\sqrt{2\pi}x}$ and symmetry.) Therefore by Borel-Cantelli, since $\sum_n \mathbb{P}(|X_n| > n) < \infty$, one has

$$\mathbb{P}(|X_n| > n \text{ infinitely often}) = 0.$$

So we can dominate this sum by

$$\sum_{n=1}^\infty nx^n,$$

which converges for any $x \in (-1, 1)$. To find the limit, we compute the characteristic function. Let Y_n be the partial sum to term n and note that since $Y_n \to Y$ almost surely, also this convergence occurs in distribution, and therefore the characteristic function of Y_n converges pointwise to that of Y. Therefore

$$\phi_Y(t) = \mathbb{E}e^{itY} = \prod_{n=1}^\infty \mathbb{E}e^{itx^nX_n}.$$

Use the fact that the characteristic function for a standard Gaussian is $e^{-t^2/2}$ to obtain

$$\prod_{n=1}^\infty e^{-(tx^n)^2/2} = \exp\left(-(t^2/2)\sum_{n=1}^\infty x^{2n}\right) = \exp\left(-\frac{t^2\frac{x^2}{1-x^2}}{2}\right).$$

This is the characteristic function of a Gaussian with mean zero and variance $x^2/(1-x^2)$.

7. Let $X_1, X_2, \ldots$ be independent random variables such that X_n has Binomial(n, p_n) distribution, for some $p_n > 0$. Show that if $np_n(1-p_n) \to \infty$, then

$$\frac{X_n - np_n}{\sqrt{np_n(1-p_n)}} \Rightarrow N(0, 1).$$

Solution: For $k \geq 1$, let $Y_{k,1}, \ldots, Y_{k,k}$ be i.i.d. Bernoulli random variables with parameter p_n. Then $Y_n = Y_{n,1} + \cdots + Y_{n,n}$ has the same distribution as X_n, has mean np_n, and has variance

$$\text{Var } X_n = \text{Var } (Y_{n,1} + \cdots + Y_{n,n}) = n\text{Var } Y_{n,1} = np_n(1 - p_n).$$

Thus the problem is asking us to show that

$$\frac{Y_n - \mathbb{E}Y_n}{\sqrt{\text{Var } Y_n}} \Rightarrow N(0, 1).$$

This will follow immediately once we show that hypothesis of Lindeberg's CLT hold. That is, we must show that if $s_n = \sqrt{\text{Var } Y_n}$, then for each $\epsilon > 0$,

$$\frac{1}{s_n^2} \sum_{i=1}^{n} \mathbb{E}Y_{n,i}^2 \mathbf{1}_{\{|Y_{n,i}| \geq \epsilon s_n\}} \to 0 \text{ as } n \to \infty.$$

As $np_n(1 - p_n) \to \infty$, one has $s_n \to \infty$. So for all large n, $\epsilon s_n > 1$. Since the $Y_{n,i}$'s are bounded in absolute value by 1, the indicator function is 0 for all large n. Thus the above expression is 0 for all large n and we are done.

8. A sequence of events $A_1, A_2, \ldots$ is said to be 1-dependent if for every $k \geq 1$, the sigma-algebras $\sigma(A_1, \ldots, A_k)$ and $\sigma(A_{k+2}, A_{k+3}, \ldots)$ are independent. Prove that if $A_1, A_2, \ldots$ are 1-dependent and E is a tail event:

$$E \in \cap_n \sigma(A_n, A_{n+1}, \ldots),$$

then $\mathbb{P}(E) = 0$ or 1.

Solution: Here the proof is almost exactly the same as that of Kolmogorov's 0/1 law. So we will use some of the tools from that proof. We aim to show that E is independent of itself, so $\mathbb{P}(E) = \mathbb{P}(E)^2$ and the result will follow.

As in the proof of the Kolmogorov 0/1 law, if we define the collection

$$\mathcal{C}_E = \{A : \mathbb{P}(A \cap E) = \mathbb{P}(A)\mathbb{P}(E)\},$$

then $\mathcal{C}_E$ is a λ-system. We claim that it contains the π-system

$$\Pi = \cup_n \sigma(A_1, A_2, \ldots, A_n).$$

Indeed, if $A \in \Pi$, then $A \in \sigma(A_1, \ldots, A_n)$ for some n. Since E is a tail event,

$$E \in \cap_k \sigma(A_k, A_{k+1}, \ldots) \subset \sigma(A_{n+2}, A_{n+3}, \ldots).$$

By assumption, $\sigma(A_1, A_2, \ldots, A_n)$ is independent of $\sigma(A_{n+2}, \ldots)$, and so E is independent of A, giving $\Pi \subset \mathcal{C}_E$.

By the $\pi - \lambda$ theorem, one has $\mathcal{C}_E \supset \sigma(\Pi)$. However

$$\sigma(\Pi) = \sigma(\cup_n \sigma(A_1, \ldots, A_n)) = \sigma(A_1, A_2, \ldots),$$

and this last sigma-algebra contains E. Therefore $E \in \mathcal{C}_E$ and E is independent of itself.